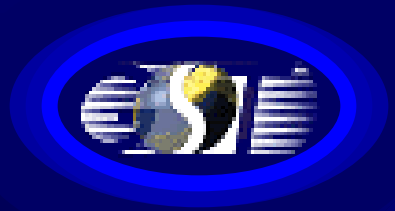
# ECHO
# EOSDIS ClearingHOuse

George Percivall
NASA/GST



From a presentation by:
Keith Wichmann – Global Science and Technology, Inc.
and Robin Pfister - NASA/GSFC

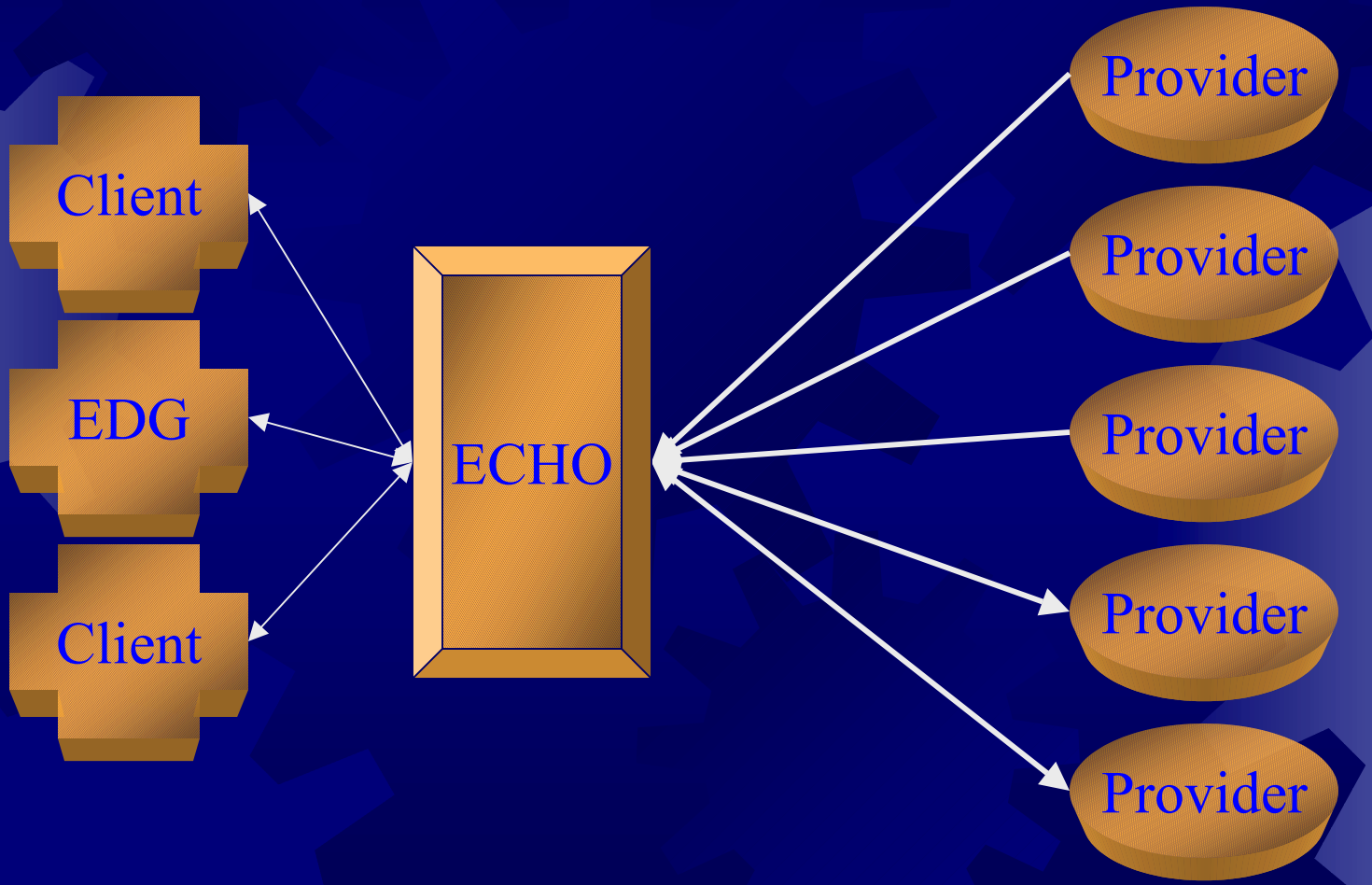
# What Is ECHO?

- Functionally
  - Clearinghouse of spatial and temporal metadata
  - Order broker
  - User and provider account service
  - Services clearinghouse and broker (future)
- Objectives
  - ECHO is a portal to Earth Science metadata.
  - It allows providers of data to share their metadata and offload some of their search responsibilities.
  - It also brokers orders from clients to the appropriate providers, providing tracking services for both the client and the provider.
  - ECHO presents a messaging interface based on XML, but does not currently provide a GUI.

# Why ECHO?

- ECHO relieves providers from providing continuous, scalable searches and order interfaces
  - Providers can schedule regular maintenance (and deal with unscheduled maintenance) without directly impacting users
  - ECHO acts as a persistent cache of metadata similar to search engines on the internet
  - ECHO brokers orders and will hold the order until the provider is ready to receive it
- ECHO allows a variety of clients to connect to it
  - ECHO provides a single place on the internet to perform geospatially and temporally enabled searches for NASA's (and possibly other's) Earth Science data.
  - ECHO applies e-commerce and B2B (business to business) technologies to support client remote access
- ECHO will also handle services
  - A clearinghouse of services will also be maintained
  - The brokering system will also handle services

# ECHO Metadata Access



# ECHO System Drivers

- Present an API for organizations to connect their own user interfaces and programs to
- Make it easy for providers of Earth Science data and services to participate in the system
- Provide searches that respond quickly
- Broker orders for both data and services
- Minimize operational costs
- Build upon advances in industry and use e-commerce systems as a model
- Build a system that can be scaled up to handle large numbers of requests

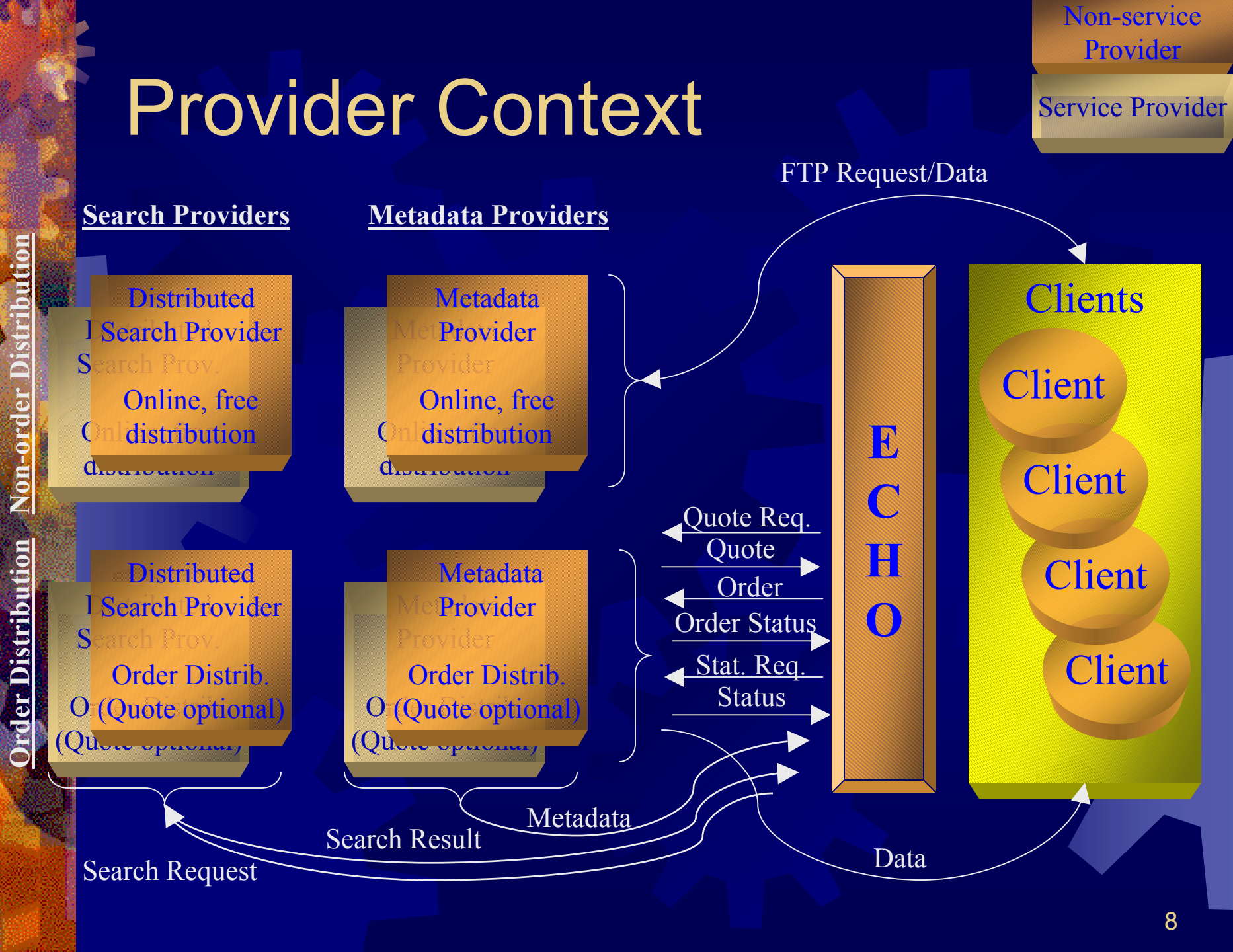
# What Does ECHO Enable?

- For Users
  - Access to a cache of Earth Science metadata in one place, regardless of provider down time
  - Order data from a single system: ECHO handles the complexity of breaking up the order for each provider.
  - GUIs that interact directly with this clearinghouse of data to find data and services and to place orders for that data can be community oriented
  - Different access paradigms can be provided (browse vs. search)
  - The service infrastructure provides a broader range of interoperable functions than would be provided by a truly centralized system
  - Truly spatial searches – not just bounding box
  - Decentralizes development of end user functionality.
- For Providers
  - Reduces the machine workload required for supporting searches, but still maintains control of data
  - Relieves scalability and availability requirements
  - Makes services available to larger community
  - Reduces need for V0 servers

# ECHO Provider Participation

- Categorization of providers by location of metadata
  - Metadata Providers
    - Participate by placing a copy of their metadata in the clearinghouse
  - Search Providers
    - Participate by receiving distributed searches and responding with metadata (Not yet implemented)
- Categorization of providers by order interactions
  - Order Distribution Providers
    - Because of the need to stage data dynamically or the need to charge for data access, an order process is needed
  - Non-order Distribution Providers
    - No order process needed, data is available via an URL
- Categorization of providers by what they provide
  - Data Providers – provide Earth Science data
  - Service Providers – provide additional processing capabilities (Not yet implemented)

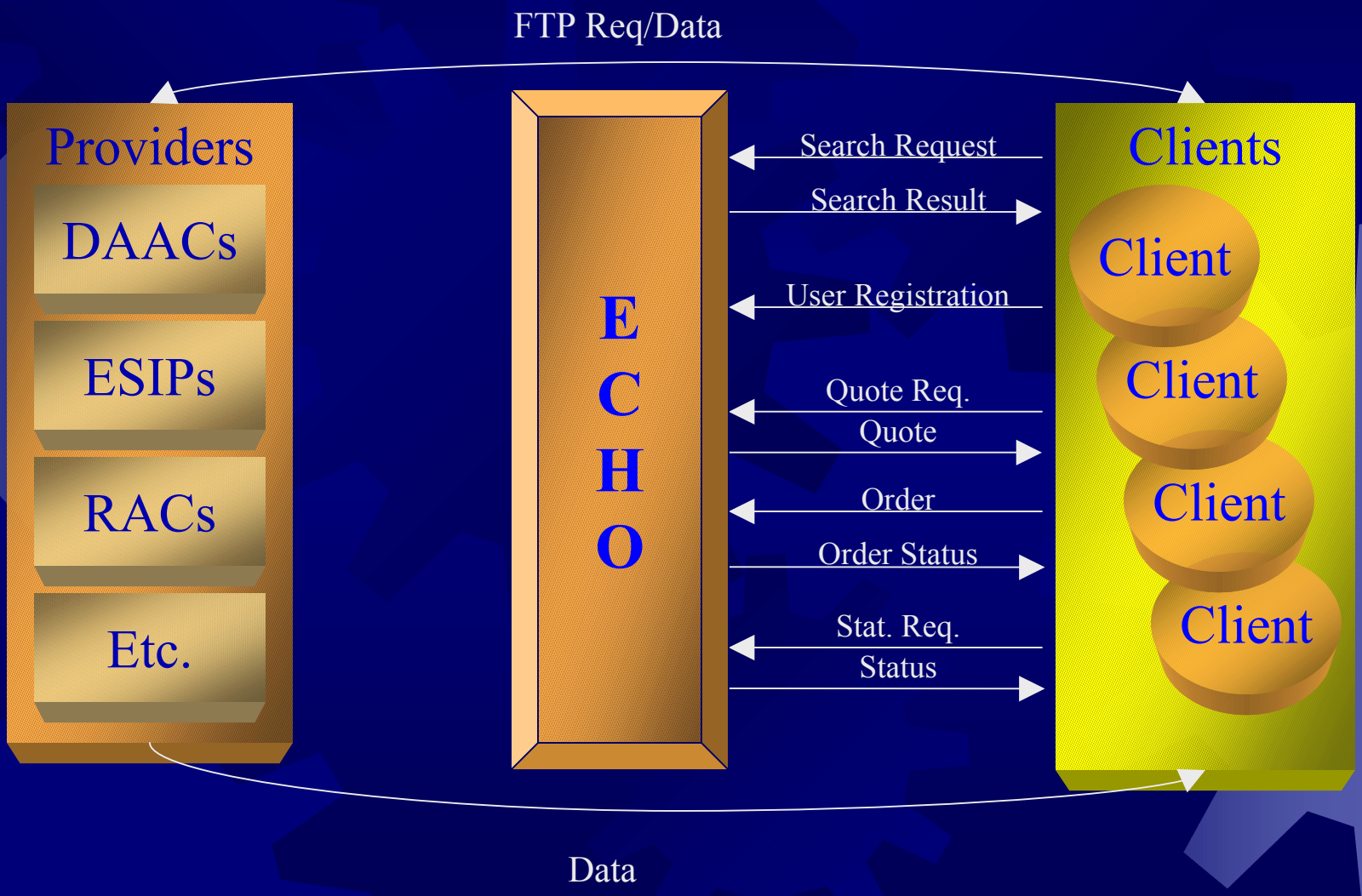
# Provider Context



# ECHO Provider Interfaces

- Provider → ECHO
  - Provider registration interface
    - Update addresses, email, etc.
  - Metadata update
    - XML files representing collections and granules sent to ECHO via FTP
  - Provider policy interface (future)
    - Method to establish parameters of how the provider participates in the ECHO system
  - User registration information interface (future)
  - Order history interface (future)
    - List of open orders, history of orders placed
- ECHO → Provider
  - Order entry, status, cancellation (f) and quote (f)

# Client Context



# ECHO Client Interfaces

- User Account Registration
  - Ability to create accounts and modify information
- Catalog Services
  - Based on OpenGIS (not fully compatible yet)
  - Ability to query metadata clearinghouse and save queries
  - Ability to manipulate result sets
  - Domain level query language (AQL)
- Order Services
  - Ability to request data, quotes, status of orders, order history

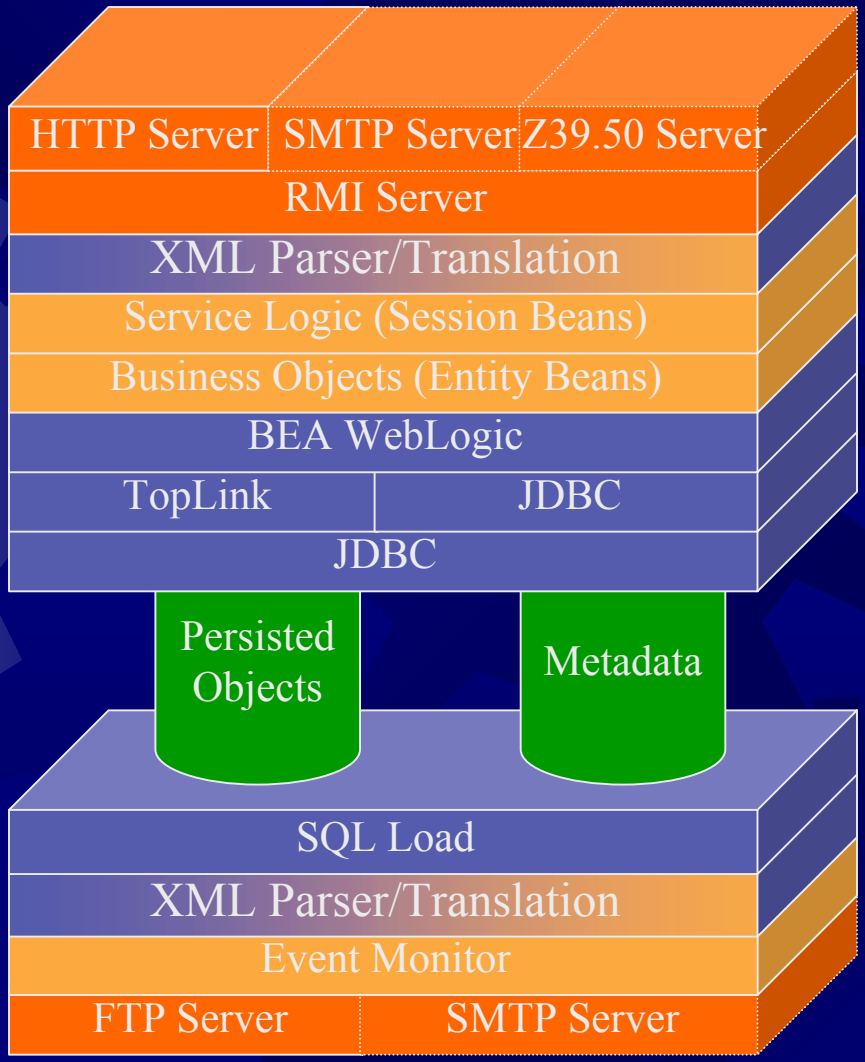
# XML

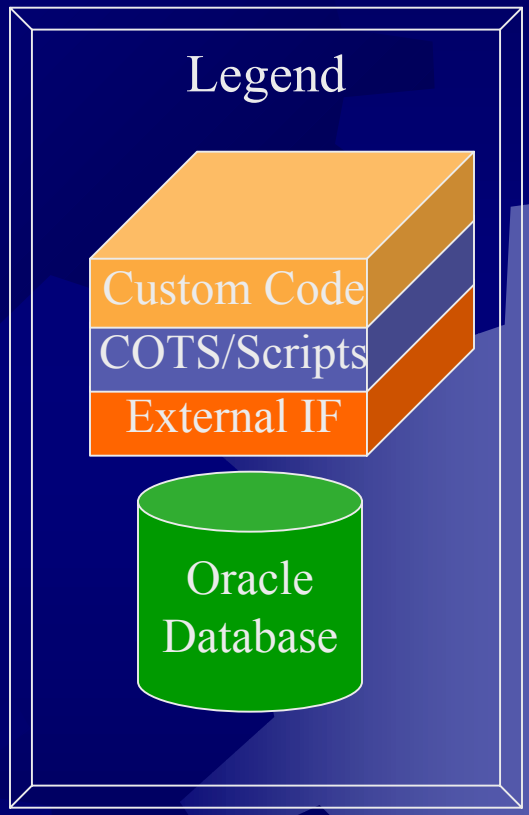
- ECHO uses XML as the basis of message and data interchange with both clients and providers
  - XML/Java mapping file allows a configuration tool to define how a conversion from the XML message to native Java types is performed.
- Synchronicity – tool to define XML based APIs
  - Generates HTML documentation, DTDs for the API, deployable EJB stubs, java parameter classes and an XML/java mapping file synchronized to each other.
- XML2XML tool
  - This tool accepts as input either XML or DTD files of two different schemas. It is then used to generate a mapping from one schema to the other. It generates an XSLT file that can then be executed to do the translation.

# Enterprise Java Beans

- BEA WebLogic is the infrastructure providing the basis for the API interfaces and the supporting business logic
  - This is the same infrastructure used at major e-commerce sites such as Amazon and British airways
- Toplink is used to connect the EJBs to oracle for persistent storage capabilities
- Java remote method invocation is used as the protocol and XML as the payload format for distributed access to the system
  - Other protocols are possible, such as HTTP, SMTP, FTP, CORBA, COM

# Internal Architecture



# ECHO Timeline

- Operational prototype software is currently up along with reference documentation (http://beamish.gsfc.nasa.gov:4000)
- By June, we expect to have a few ECS datasets in the system and available for order
- ORNL data is already available
- New functions and improvements should roll out about every 4 months
- Project website: http://dangermouse.gst.com/ECHO